Proceedings of ICEIPI 2026 Symposium: Inclusive Education, Cultural Transformation, and the Ethical Dimensions of Learning
DOI: 10.54254/2753-7048/2026.LD36055

Differential Compensatory Effects of AR-Enhanced STEAM Instruction on Seventh Grad' Spatial Ability: A Quasi-Experimental Study

Mengqi Wang*, Xinlu Wang

School of Faculty of Education, Shandong Normal University, Jinan, China
**Corresponding Author. Email: 16678901818@163.com*

Abstract. Gender and proficiency gaps in spatial ability are prevalent among junior high school students learning geometry, while traditional physical teaching tools often provide limited support for students with lower spatial abilities. This quasi-experimental study investigated whether GeoGebra AR-based STEAM lessons could compensate for spatial deficits and enhance students' geometric reasoning. A total of 62 seventh-grade students participated in an eight-session intervention, with students assigned to either an AR experimental group or a physical manipulative control group. Spatial ability was evaluated using the Mental Rotation Test (MRT), Mental Paper Folding Test, and Spatial Visualization Ability Test (SBST), with both accuracy and response time analyzed. The results indicated that both instructional approaches improved students' spatial performance; however, only the AR intervention effectively reduced gender disparities and produced substantial improvements among low-ability learners, particularly in abstract mental rotation tasks. Furthermore, the compensatory effects of AR extended to hands-on STEAM design and creation activities. These findings provide empirical evidence supporting the integration of AR technologies into junior high 3D geometry education.

Keywords: Augmented reality, spatial ability, STEAM education, Differential Compensation

1. Introduction

Spatial ability is critical for junior high solid geometry and STEAM design, encompassing mental rotation, paper folding and solid sectioning [1, 2]. Early adolescents displaymarked spatial disparities: females lag behind males on abstract spatial tasks, and learners are grouped into high/medium/low ability tiers based on cognitive differences [3]. Low-ability struggle with 2D-3D mental transformation. Traditional static physical models cannot dynamically rotate or dissect geometric shapes, failing to scaffold low-ability learners, thereby widening spatial gaps [4]. AR visualizes abstract spatial reasoning and reduces cognitive load, yet most existing AR studies focus on university engineering students and lacksystematic quasi-experiments targeting junior-high geometry beginners and analyzing both accuracy and response time [2].

This study adopted a pre-post quasi-experiment to investigate AR's differentiated compensation effect, with four research questions: RQ1: Can both teaching modes improve seventh graders' spatial

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2753-7048/2026.LD36055

ability? RQ2: Can AR reduce gender gaps in spatial performance? RQ3: Does AR specially remediate low-ability students' spatial deficits? RQ4: Is AR's promotion stronger for more abstract spatial tasks, and how does response time relate to accuracy?

2. Related work

Three frameworks underpin this research: Multiple Intelligences Theory, Cognitive Load Theory and Constructivist Learning Theory. Gardner's theory suggests spatial intelligence is developable through external learning experiences [5]. Cognitive Load Theory notes abstract spatial imagery consumes substantial mental resources; without dynamic visual aids, low-ability students face severe cognitive overload [6]. Constructivism argues learners construct spatial cognition through concrete interactive materials, and AR visualization supports autonomous knowledge construction [7]. Existing studies suggest digital mixed-reality tools externalize internal spatial reasoning and reduce cognitive burden [8]. Compared with paper-based learning, interactive digital environments extend student response time, and prolonged exploratory observation correlates positively with test accuracy [9]. Prior subgroup analyses indicate digital visualization delivers far stronger improvements for low-spatial-ability learners than for advanced students [10].

H1: Both experimental and control groups show significant post-test spatial ability growth.

H2: The AR experimental group achieves larger overall spatial gains than the control group.

H3: Low-ability students in the AR group gain significantly more than control low achievers; AR reduces intra-class spatial stratification.

H4: Females' spatial improvement rate lags behind males under traditional instruction, while AR eliminates this gender-based disparity.

3. Methodology

3.1. Participants

A total of 62 seventh-grade students from two parallel classes in a middle school participated in this study. All participants were between 12 and 13 years old and had not received formal instruction in solid geometry prior to the experiment. A quasi-experimental design was adopted, with one class assigned to the experimental group (AR-enhanced STEAM instruction) and the other serving as the control group (traditional manipulative-based STEAM instruction). Before the intervention, all students completed a spatial ability pre-test to establish equivalence between groups. The same teacher delivered all lessons, and both groups received identical instructional content, class hours, assignments, and STEAM project tasks. The only difference was the instructional tool used during inquiry activities. Gender information was collected and used for subsequent analyses examining the differential effects of AR instruction.

3.2. Instruments

The experimental group used GeoGebra 3D Calculator with integrated AR functionality. The software enabled students to visualize and manipulate three-dimensional geometric objects through interactive operations such as rotation, folding and unfolding, and cross-section exploration. The control group completed the same learning activities using traditional physical manipulatives, including solid models, cardstock, and plasticine. No AR software or digital three-dimensional visualization tools were used in the control condition.

DOI: 10.54254/2753-7048/2026.LD36055

Three standardized spatial cognition tests were administered using Inquisit 6 software, which together constructed a three-dimensional spatial ability assessment system across three independent dimensions: mental rotation, planar-to-solid folding, and geometric solid sectioning. The software automatically recorded quantitative data including accuracy (percentage correct) and response time. Based on the original norms of the classic tests and localized empirical data from Chinese adolescents, the statistical criterion of mean±one standard deviation was adopted to classify participants into three spatial ability levels: high, medium, and low.

3.3. Curriculum design

An 8-lesson introductory solid geometry curriculum integrating STEAM was designed, with each lesson lasting 40 minutes. The curriculum was structured around three core ability objectives: mental rotation, solid folding, and geometric solid sectioning. It integrated interdisciplinary elements from science (S), technology (T), engineering (E), art (A), and mathematics (M). The experimental group and control group shared identical curriculum objectives, knowledge points, and STEAM project tasks; the only difference lay in the hands-on medium.

3.4. Experimental procedure

The study comprised three phases: a pre-test, an intervention, and a post-test. The pre-test, conducted one day before instruction, included three spatial ability tests.Based on these results, students were classified into high, medium, and low ability levels; gender was also recorded, and low-ability students were designated as the primary subgroup to examine compensatory effects. The intervention lasted two consecutive weeks with a total of eight lessons, during which the two groups differed only in instructional medium while the teacher, class hours, exercises, and after-class projects were kept identical. The post-test was administered on the day after the final lesson, and the resulting data were exported for statistical analysis.

4. Result

4.1. Overall improvement and group differences

Pre-test comparisons indicated that the experimental AR group and the traditional control group were equivalent on all spatial measures, with all *p* values exceeding .05. Baseline data also revealed a clear gender gap: male students scored higher and responded faster than female students. The Mental Rotation Test, or MRT, proved the most difficult, yielding the lowest accuracy and the longest response time. Based on pre-test total scores, students were divided into high, medium and low ability levels.

Table 1. Pre-test, post-test scores and improvement by group and test type

Group	Test Type	Pre-test (M ± SD)	Post-test (M ± SD)	Improvement (M ± SD)	t(paired)	p	Cohen's d
Experimental	MRT	52.3 ± 10.2	78.6 ± 9.4	26.3 ± 8.1	8.45	<.001	1.52
	Mental Paper Folding	60.5 ± 9.8	82.1 ± 7.5	21.6 ± 7.3	7.23	<.001	1.31
	SBST	57.8 ± 10.5	79.4 ± 8.2	21.6 ± 7.6	6.98	<.001	1.26

DOI: 10.54254/2753-7048/2026.LD36055

Table 1. (continued)

	MRT	53.1 ± 11.0	68.2 ± 10.1	15.1 ± 7.2	4.92	<.001	0.89
Control	Mental Paper Folding	61.2 ± 8.9	74.3 ± 8.6	13.1 ± 6.5	4.56	<.001	0.82
	SBST	58.3 ± 9.7	71.6 ± 9.0	13.3 ± 6.8	4.78	<.001	0.86

After the eight-lesson intervention, both groups improved significantly on every subtest, as confirmed by paired-samples *t*-tests, all with *p* < .001. However, the AR group showed substantially larger gains. As summarised in Table 1, the experimental group's improvement on MRT was 26.3 percentage points, rising from 52.3 to 78.6, whereas the control group improved only 15.1 points. On Mental Paper Folding, the AR group gained 21.6 points versus 13.1 in the control group; on the Santa Barbara Solids Test, the gains were 21.6 and 13.3 points respectively. Independent-samples *t*-tests indicated that the experimental group's improvements were significantly greater on all three tests: for MRT, *t* = 5.12 with *p* < .001; for Paper Folding, *t* = 4.01 with *p* < .001; and for SBST, *t* = 3.89 with *p* < .001. Effect sizes, measured by Cohen's *d*, were large for the AR group, ranging from 1.26 to 1.52, while the control group showed moderate effects from 0.82 to 0.89. Additionally, the AR group had longer average response times across all item types; for example, on the MRT they took 28.3 seconds , compared with19.5 seconds for the control group, a difference that was statistically significant with *t* = 5.67 and *p* < .001. This longer time reflects active, exploratory engagement rather than inefficiency.

4.2. Interaction effects

4.2.1. Instructional approach × gender/ability level

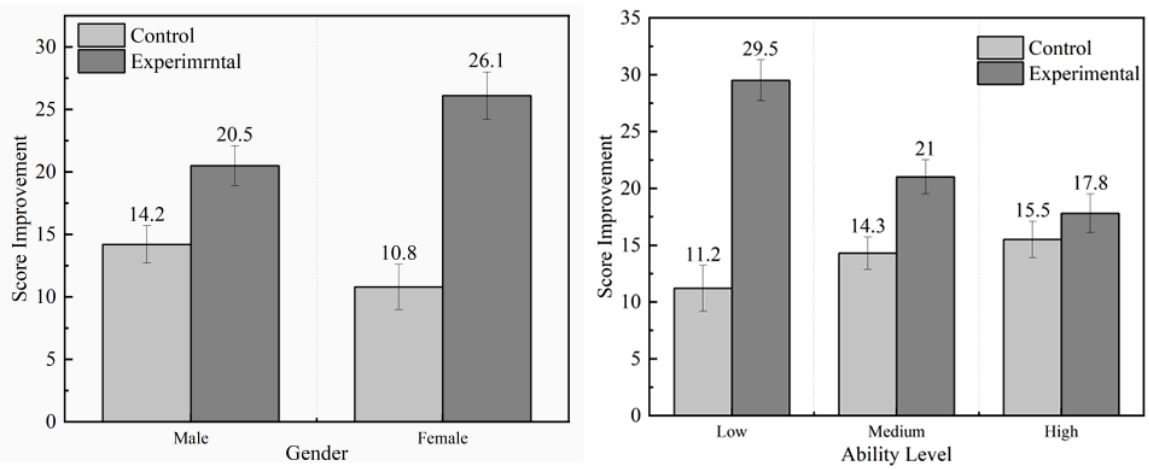


Figure 1. Interaction between instructional approach and gender/ability level on spatial ability improvement

As depicted in Figure 1, in the control group male students improved significantly more than female students did, with a mean gain difference of 8.2 percentage points and a *p* value less than.05. Thus traditional instruction did not close the gender gap. In the AR group, however, female students showed substantially larger gains, and the male-female difference was reduced to a non-significant level with *p* = .23. This outcome suggests that AR compensates for female students' spatial learning disadvantages by offering concrete visual support.

DOI: 10.54254/2753-7048/2026.LD36055

Figure 1 also presents the stratified effects. High-ability students showed almost no difference between the two conditions, with a gain gap of less than 3 percentage points. Medium-ability students benefited slightly more from AR, showing a difference of about 5 points. The most striking result came from low-ability students: those in the AR group improved far more than their control counterparts, with a gain difference of 18.4 percentage points, which was highly significant (*p* < .001, *d* = 1.68). This indicates that AR serves as a cognitive scaffold that benefits students with the weakest initial spatial ability the most.

4.2.2. Instructional approach × test type

The largest between-group difference in improvement occurred for the MRT, where the AR group outperformed the control group by 11.2 percentage points. For the Paper Folding Test the difference was 8.5 points, and for the SBST it was 8.3 points. This pattern suggests that the more abstract the spatial task, the greater the compensatory advantage of AR, because mental rotation places the heaviest demand on internal visualisation, a process that AR externalises effectively.

4.3. Three-way interaction and correlational findings

A three-way ANOVA with group, gender and ability level as factors was performed on total gain scores. The analysis revealed a significant main effect of group, with $F(1,56) = 20.12$, *p* < .001, $\eta^2 = 0.26$, and a significant main effect of ability level, $F(2,56) = 15.34$, *p* < .001, $\eta^2 = 0.34$. The Group × Gender interaction was significant, $F(1,56) = 4.15$, *p* = .046, $\eta^2 = 0.07$, as was the Group × Ability Level interaction, $F(2,56) = 8.92$, *p* < .001, $\eta^2 = 0.22$. The three-way interaction did not reach significance, with *p* = .084. Post-hoc simple effect analyses identified low-ability female students in the AR group as the subgroup showing the largest overall improvement, suggesting that AR simultaneously reduced both gender-based and ability-based disadvantages.

Pearson correlations between post-test response time and accuracy were positive and significant for all subtests. In the AR group, the correlation for MRT was 0.52, for the Paper Folding Test 0.44, and for the SBST 0.46; in the control group, the corresponding coefficients were 0.48, 0.39 and 0.41, all with *p* < .05. These results support that thoughtful, deliberate observation is associated with better performance, and the AR group's longer response times played a role in their superior accuracy.

4.4. Complementary STEAM project performance

Throughout the eight lessons, students completed hands-on STEAM tasks such as designing packaging boxes, sectioning solids and creating three-dimensional art pieces. In the control group, low-ability students often produced dimensionally inaccurate or structurally unstable products because they could not mentally visualise the shapes beforehand. In contrast, the AR group students used the software to preview unfolding nets and cross-sections before physical construction, which led to better engineering precision and artistic quality. This suggests that AR's compensatory effect extends beyond standardised tests into authentic, interdisciplinary practice.

DOI: 10.54254/2753-7048/2026.LD36055

5. Discussion

5.1. Differential mechanisms of instructional mediums affecting spatial cognitive development

Both traditional hands-on instruction (wooden blocks, cardstock, etc.) and GeoGebra AR-enhanced teaching effectively improve seventh-graders' 3D spatial ability, through distinct cognitive pathways. Traditional manipulation relies on tactile feedback and concrete operations, an approach that aligns with the "concrete-to-abstract" principle for younger students. However, physical manipulatives have inherent limitations—fixed shapes, non-instant deformations, and irreversible sectioning—forcing teachers to frequently replace materials when demonstrating different viewpoints or oblique sections. This results in fragmented acquisition of spatial features and discontinuous cognitive construction. In contrast, AR overcomes physical constraints by generating dynamic geometric solids in mixed reality, allowing real-time multi-angle rotation, unfolding, and on-demand sectioning with a single click. Continuous visualization of structural changes allows students to systematically observe multiple configurations within a short time, thereby forming a coherent spatial representation. The AR group showed longer response times but significantly higher accuracy; this is not inefficiency but reflects deep observation and autonomous reasoning fostered by the visualized environment. Sufficient visual inquiry enables internalization of spatial logic, which is the core cognitive reason for the AR group's superior spatial improvement.

5.2. How AR eliminates gender-based spatial disparities

Baseline data showed existing gender differences among 12- to 13-year-olds without formal geometry training: males outperformed females in mental rotation and cross-section reasoning. This stems from early adolescent differences in spatial representation modes—males tend to use mental rotation implicitly, while females rely more on external concrete cues and are prone to cognitive blockage when such cues are absent. Under traditional instruction, limited manipulatives cannot provide continuous dynamic references, forcing female students into unsupported abstract visualisation, which amplifies their initial disadvantage. AR precisely supplies the concrete external support that females need, and transforms mental rotation and virtual sectioning into manipulable real-time visual scenarios. By externalising implicit cognitive tasks, AR reduces the cognitive load for female students. Thus, AR does not alter biological spatial aptitude; rather, itbridges the learning gap by optimising the external environment to match female students' cognitive processing styles, and compensates for their spatial learning disadvantages.

5.3. Cognitive interpretation of AR's stratified compensatory effects

Stratified analyses show that high- and medium-ability students exhibited only marginal differences between the two instructional conditions, whereas low-ability students benefited most from AR. This is rooted in pre-existing spatial cognitive reserves. High-ability students have already built a basic spatial representation system and can perform 3D-to-2D transformations with physical objects; therefore, traditional manipulatives suffice for their development. Medium-ability students possess partial representational skills and can gradually fill gaps through trial-and-error with hands-on materials, so the gain difference is limited. Low-ability students, however, lack initial spatial representation and cannot mentally manipulate 3D structures; trial-and-error with physical materials is extremely inefficient, and repeated failures often induce cognitive fatigue, hindering their progress. AR acts as a cognitive scaffold through dynamic visualisation, progressively guiding

DOI: 10.54254/2753-7048/2026.LD36055

low-ability students to establish logical relationships among solids, nets, and cross-sections, and thereby filling cognitive gaps step by step. Sub-test findings further clarify: mental rotation (MRT) is the most abstract and demands the strongest internal representation, so low-ability students gained most on this dimension. For paper-folding and solid sectioning, physical manipulation (e.g., folding cardstock, cutting clay) can partially meet learning needs, hence the improvement gap is smaller. This pattern reveals that the more abstract the spatial task and the more deficient the student's initial reserve, the more pronounced AR's compensatory value.

5.4. Deepening AR's pedagogical value within STEAM's interdisciplinary framework

Our curriculum integrated hands-on tasks across the five STEAM dimensions (Science, Technology, Engineering, Art, Mathematics), thereby extending AR's compensatory effects from paper-and-pencil tests to authentic practice. In science, AR intuitively visualises geometric structures and cutting rules, helping students understand the construction of everyday objects. In engineering, AR assists in planning product dimensions in advance, reducing rework. In art, rich 3D materials broaden modelling ideas. This demonstrates that AR's compensatory effects is interdisciplinary and aligns perfectly with the STEAM learning context of this study.

6. Conclusion

Synthesising the experimental results and initial hypotheses, both traditional manipulative-based instruction and AR-assisted instruction proved effective in improving seventh-grade spatial ability. Compared with traditional instruction, AR visualisation not only narrowed the observed gender gap but also provided targeted support for low-ability students, thereby alleviating within-class ability disparities. The improvement was most pronounced on the most abstract task (MRT), and this compensatory advantage extended equally to hands-on STEAM activities such as engineering construction and artistic creation.

Constrained by research conditions, this study only selected two classes for an eight-lesson, short-term experiment, with a limited sample, a single AR software tool, and primarily reliance on standardised tests, thus limiting generalisability. Future research should expand the sample to include different schools and grade levels, extend the intervention period, compare multiple AR tools, and enrich the evaluation system with hands-on product scoring and classroom observation. Further studies should continue to investigate differentiated developmental effects on students with various characteristics in AR-enhanced classrooms.

References

[1] Carbonell Carrera, C. J., Jaeger, A., Saorín, J. L., Melián, D., & Cantero, J. (2021). Minecraft as a block building approach for developing spatial skills. *Entertainment Computing, 38*, 100427. https://doi.org/10.1016/j.entcom.2021.100427

[2] Li, Y., Yang, Y., Yao, Z., & Xu, G. (2020). Virtual 3D environment for exploring the spatial ability of students. *Virtual Reality & Intelligent Hardware, 2*(6), 556-568. https://doi.org/10.1016/j.vrih.2020.08.001

[3] Wang, L. (2020). Mediation relationships among gender, spatial ability, math anxiety, and math achievement. *Educational Psychology Review, 32*(3), 1-15. https://doi.org/10.1007/s10648-019-09487-z

[4] Sari, D. P., Syahputra, E., & Surya, E. (2018). An analysis of spatial ability and self-efficacy of students in cooperative learning by using jigsaw at smas muhammadiyah 8 kisaran. *American Journal of Educational Research, 6*(8), 1238-1244. https://doi.org/10.12691/education-6-8-25

[5] Wang, D. (2025). Application and significance of multiple intelligences theory in general educational development planning. *Journal of Higher Education Research, 6*(4).

DOI: 10.54254/2753-7048/2026.LD36055

[6] Kenneally, C. D., & Bentley, B. (2024). A cognitive load approach to molecular geometries: Augmented reality technology and visuospatial abilities in chemistry. *Education Sciences, 14*(9), 1036. https://doi.org/10.3390/educsci14091036

[7] Qi, M. (2025). Study on the effect of autonomous learning of college students in the online education model. *Journal of Innovation and Development, 10*(2), 9-11.

[8] Al-Ansi, A. M., Jaboob, M., Garad, A., & Al-Ansi, A. (2023). Analyzing augmented reality (AR) and virtual reality (VR) recent development in education. *Social Sciences & Humanities Open, 8*(1), 100532. https://doi.org/10.1016/j.ssaho.2023.100532

[9] Liu, C., Zhang, D., Pandolpho, M., & Yi, J. (2023). Effects of principled digitalized interactive components on geometry computer-based assessment in high schoolers with learning difficulties in mathematics: A preliminary investigation. *Assessment for Effective Intervention, 49*(1), 29-40. https://doi.org/10.1177/1534508423120323

[10] Lee, E. A. L., & Wong, K. W. (2014). Learning with desktop virtual reality: Low spatial ability learners are more positively affected. *Computers & Education, 79*, 49-58. https://doi.org/10.1016/j.compedu.2014.07.010